

## **IN THE CLAIMS**

Set forth below is a clean set of the amended claims. Such amendments are shown in the attachment.

Please amend claims 1, 16 and 17 as follows:

1. A system for interactive language instruction comprising:
  - a first module configured to convert input text to audible speech in a selected language, the audible speech being patterned after a model;
  - a user interface configured to receive utterances spoken by a user in response to a prompt to replicate the audible speech; and,
  - a second module configured to recognize the utterances and provide feedback to the user, the feedback being comprised of a confidence measure reflecting a precision at which the user replicates each portion of the audible speech in the selected language based on a comparison of the utterances to one of the audible speech and the model.
  
16. A system comprising:
  - a first module configured to convert input text to audible speech in a selected language, the audible speech indicative of a model, wherein the model is one of a predictive model, a phoneme model, a diphone model, and a dynamically generated model;
  - a second module synchronized to the first module, the second module producing an animated image of a human face and head pronouncing the audible speech;
  - a user interface positioned to receive utterances spoken by a user in response to a prompt to replicate the audible speech; and,
  - a third module configured to recognize the utterances and provide feedback to the user, the feedback being comprised of at least one of a score, an icon and an audio segment reflecting a precision at which the user replicates each portion of the speech in the selected language based on a comparison of the utterances to one of the audible speech and the model.

17. A method for voice interactive language instruction comprising:  
converting input text data to audible speech data;  
generating audible speech comprising phonemes based on the audible speech data;  
outputting the audible speech through an audio output device;  
generating an animated image of a face and head pronouncing the audible speech;  
synchronizing the audible speech and the video image;  
prompting a user to replicate the audible speech;  
recognizing utterances generated by the user in response to the prompting;  
comparing the audible speech to the utterances; and,  
providing feedback to the user based on the comparison, the feedback comprised of at least one of a score, an icon and an audio segment reflecting a precision at which the user replicates each portion of the audible speech.